

=> d his

(FILE 'HOME' ENTERED AT 14:14:12 ON 02 APR 2003)

FILE 'CAPLUS' ENTERED AT 14:14:18 ON 02 APR 2003

L1 530 S DATABASE? (6A) (FINGERPRINT? OR PATTERN?)

L2 21 S L1 AND RATIO?

L3 53 S L1 AND (TAG? OR MARKER?)

L3 ANSWER 40 OF 53 CAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1999:550714 CAPLUS

DOCUMENT NUMBER: 131:267927

TITLE: Motifer, a search tool for finding amino acid sequence patterns from nucleotide sequence databases

AUTHOR(S): Jornvall, Henrik

CORPORATE SOURCE: Laboratory of Molecular Neurobiology, Department of Neuroscience, Karolinska Institutet, Stockholm, S-171 77, Swed.

SOURCE: FEBS Letters (1999), 456(1), 85-88

CODEN: FEBLAL; ISSN: 0014-5793

PUBLISHER: Elsevier Science B.V.

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Motifer is a software tool able to find directly in nucleotide databases very distant homologues to an amino acid query sequence. It focuses searches on a specific amino acid pattern, scoring the matching and intervening residues as specified by the user. The program has been developed for searching databases of expressed sequence tags (ESTs), but it is also well suited to search genomic sequences. The query sequence can be a variable pattern with alternative amino acids or gaps and the sequences searched can contain introns or sequencing errors with accompanying frame shifts. Other features include options to generate a searchable output, set the maximal sequencing error frequency, limit searches to given species, or exclude already known matches. Motifer can find sequence homologs that other search algorithms would deem unrelated or would not find because of sequencing errors or a too large no. of other homologs. The ability of Motifer to find relatives to a given sequence is exemplified by searches for members of the transforming growth factor- $\beta$  family and for proteins contg. a WW-domain. The functions aimed at enhancing EST searches are illustrated by the 'in silico' cloning of a novel cytochrome P 450 enzyme.

REFERENCE COUNT: 11 THERE ARE 11 CITED REFERENCES

L3 ANSWER 41 OF 53 CAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1999:546667 CAPLUS

DOCUMENT NUMBER: 131:322182

TITLE: SMILES extensions for pattern matching and molecular transformations: applications in chemoinformatics

AUTHOR(S): Bone, Richard G. A.; Firth, Michael A.; Sykes, Richard A.

CORPORATE SOURCE: Proteus Molecular Design Ltd., Macclesfield Cheshire, SK11 0JL, UK

SOURCE: Journal of Chemical Information and Computer Sciences (1999), 39(5), 846-860

CODEN: JCISD8; ISSN: 0095-2338

PUBLISHER: American Chemical Society

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The selection and modification of atoms or functional groups underly many of the manipulations central to mol. modeling. It has become even more important to automate these tasks with the current prevalence of work with large databases or mols. SUPER-SMILES, a conceptually simple set of extensions to SMILES line notation, whose key features are addn. and deletion facilities, macros, atom tagging, disjunctions, and constraints, has been developed. This superset of SMILES enables the performance of transformations on individual mol. structures or across members of a database with a pattern-matching protocol. The principal advantage of SUPER-SMILES is the ability to specify chem. reactions with a very simple augmentation of the SMILES line notation. SUPER-SMILES represents a unified approach to mol. structure specification and modification and can easily be applied to large datasets of mols. This functionality has been implemented within the PROMETHEUS suite of CAMD programs. Its use in performing such operations as atom-type assignment, protonation of mols., valency checking, and hydrogen addn., is demonstrated.

REFERENCE COUNT: 37 THERE ARE 37 CITED REFERENCES

L3 ANSWER 42 OF 53 CAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1999:505479 CAPLUS

DOCUMENT NUMBER: 132:90246

TITLE: Rapid making a photoreceptor protein catalogue using mass spectrometry

AUTHOR(S): Nishizawa, Yuji; Matsumoto, Hiroyuki

CORPORATE SOURCE: Department of Anatomy, School of Medicine, Nagoya University, Tsurumai-cho, Showa-ku, Nagoya, 466-8550, Japan

SOURCE: Journal of the Mass Spectrometry Society of Japan (1999), 47(3), 177-182

CODEN: JMSJEY; ISSN: 1340-8097

PUBLISHER: Nippon Shitsuryo Bunseki Gakkai

DOCUMENT TYPE: Journal

LANGUAGE: Japanese

AB To assemble a catalog of proteins expressed in mammal photoreceptor cells, following exptl. strategy was carried out: the prepn. of bovine photoreceptor cell monolayer (PCL) and the methodol. for analyzing proteins on 2-D gels and mass spectrometry. Trypsin digested peptides were sepd. by HPLC and analyzed by a matrix-assisted laser

desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer. Some peptides were sequenced by Edman degrdn. To identify the proteins, each peptide-mass data set was referred to online peptide-mass fingerprinting database. To standardize the 2-D gel profile, we detd. the pI and mol. wt. coordinates of some of the major proteins in the photoreceptor cells by running internal proteins marker and by considering theor. values based on the known sequences. Fourteen major 2-D gel spots were identified and listed to our catalog.

L3 ANSWER 43 OF 53 CAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1999:96402 CAPLUS

DOCUMENT NUMBER: 130:149542

TITLE: Databases for compilation and analysis of information about gene expression

INVENTOR(S): Balaban, David J.

PATENT ASSIGNEE(S): Affymetrix, Inc., USA

SOURCE: PCT Int. Appl., 35 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 2

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9905323	A1	19990204	WO 1998-US15151	19980724
US 2002015948	A1	20020207	US 1998-20743	19980209
US 6420108	B2	20020716		
WO 9905591	A2	19990204	WO 1998-US15469	19980724
WO 9905591	A3	19990520		
EP 1009861	A1	20000621	EP 1998-936972	19980724
JP 2001511546	T2	20010814	JP 2000-504290	19980724
JP 2001511550	T2	20010814	JP 2000-504489	19980724
JP 2001511529	T2	20010814	JP 2000-504502	19980724
JP 11342000	A2	19991214	JP 1999-30642	19990208
US 2001018642	A1	20010830	US 2001-737838	20010326
US 2002150932	A1	20021017	US 2001-28748	20011221
PRIORITY APPLN. INFO.:			US 1997-53842P	P 19970725
			US 1997-69198P	P 19971211
			US 1997-69436P	P 19971211
			US 1998-20743	A 19980209
			US 1998-122304	A1 19980724
			WO 1998-US15151	W 19980724
			WO 1998-US15456	W 19980724
			WO 1998-US15469	W 19980724

AB An efficient and easy to use query system for a gene expression database is described. Using such a system, one can easily identify genes or expressed sequence tags whose expression correlates to

particular tissue types. Various tissue types may correspond to different diseases, states of disease progression, different organs, different species, etc. Researchers may now use large scale gene expression databases to full advantage.

REFERENCE COUNT: 3 THERE ARE 3 CITED REFERENCES AVAILABLE

L3 ANSWER 44 OF 53 CAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1999:6423 CAPLUS

DOCUMENT NUMBER: 130:205581

TITLE: Cell line characterization and authentication

AUTHOR(S): Kaplan, Joseph; Hukku, Bharati

CORPORATE SOURCE: Cell Culture Laboratory, Department of Pediatrics, Wayne State University School of Medicine, Children's Hospital of Michigan, Detroit, MI, 48201, USA

SOURCE: Methods in Cell Biology (1998), 57(Animal Cell Culture Methods), 203-216  
CODEN: MCBLAG; ISSN: 0091-679X

PUBLISHER: Academic Press

DOCUMENT TYPE: Journal; General Review

LANGUAGE: English

AB A review, with 35 refs. Research and development involving the use of cell lines require precise knowledge of the purity and species of origin of the cell lines used. This can only be assured by periodic monitoring of cultured cell lines for possible contamination by other cells and for characteristics that authenticate the cell line identity. In the absence of such monitoring, inter- and intraspecies cell line contaminations are likely to occur in the labs. of unsuspecting investigators and can result in the generation of mistaken conclusions with an attendant loss of investigators' time, effort, and resources. This chapter provides a history and an overview of the methods that have been developed for cell line authentication, the type of information each of these different methods provides, and how synthesis of that information can be used to characterize a cell line and confirm its identity. An effective cell line monitoring strategy is described that involves testing for a combination of genetic markers, including cell membrane species antigens, isoenzymes, chromosomes, and DNA fingerprints, and use of databases for each marker system to compare the results obtained with a test cell culture with results from an extensive panel of previously tested cell lines. (c) 1998 Academic Press.

REFERENCE COUNT: 39 THERE ARE 39 CITED REFERENCES

L3 ANSWER 46 OF 53 CAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1998:697567 CAPLUS

DOCUMENT NUMBER: 130:63133

TITLE: Microfabricated device coupled with an electrospray ionization quadrupole time-of-flight mass spectrometer: protein identifications based on enhanced-resolution mass spectrometry and tandem mass spectrometry data

AUTHOR(S): Figeys, Daniel; Lock, Chris; Taylor, Lorne; Aebersold, Ruedi  
CORPORATE SOURCE: Institute for Marine Biosciences, National Research  
Council Canada, Halifax, NS, Can.  
SOURCE: Rapid Communications in Mass Spectrometry (1998), 12(20), 1435-1444  
CODEN: RCMSEF; ISSN: 0951-4198  
PUBLISHER: John Wiley & Sons Ltd.  
DOCUMENT TYPE: Journal  
LANGUAGE: English

AB We describe the coupling of a microfabricated fluidic device to an electrospray ionization (ESI) quadrupole time-of-flight mass spectrometer (QqTOFMS) for the identification of protein samples. The microfabricated devices consisted of three reservoirs connected via channels to a main capillary, which in turn was linked via a microspray interface to the QqTOFMS. Here we present preliminary results obtained using this system. Standardized solns. of myoglobin tryptic digest were analyzed indicating a limit of detection at the low to sub fmol/pL. The combination of the microfabricated device for rapid sample delivery and the rapid acquisition capability, enhanced resoln. and mass accuracy of the QqTOF offers unique possibilities for the rapid identification of proteins by database searching. This platform can generate MS data suitable for protein database searching by the peptide-mass fingerprinting approach and MS/MS data suitable for protein database searching. Here the results of the two database-searching approaches are compared and the possibilities of combining the two approaches for rapid identification of protein are discussed. Also, we present a comparison of the results obtained using the three-position microfabricated device coupled to the ESI-QqTOFMS and to an ESI-ion trap MS. Finally the combination of C-terminal 180 labeling of peptides and the microfabricated system for automated combined peptide-mass fingerprinting and sequence-tag database searching is discussed.

REFERENCE COUNT: 42 THERE ARE 42 CITED REFERENCES

L3 ANSWER 50 OF 53 CAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1995:214419 CAPLUS

DOCUMENT NUMBER: 122:27645

TITLE: Cellular fatty acid composition as a chemotaxonomic marker for the differentiation of phenospecies and hybridization groups in the genus *Aeromonas*

AUTHOR(S): Huys, Geert; Vancanneyt, Marc; Coopman, Renata; Janssen, Paul; Falsen, Enevold; Altwegg, Martin; Kersters, Karel

CORPORATE SOURCE: Laboratorium voor Microbiologie, Universiteit Gent, Ghent, B-9000, Belg.

SOURCE: International Journal of Systematic Bacteriology (1994), 44(4), 651-8  
CODEN: IJSBA8; ISSN: 0020-7713

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Ninety genotypically characterized *Aeromonas* strains, including members of

all 14 currently established genospecies, were studied by performing gas-liq. chromatog. anal. of their cellular fatty acid Me esters (FAMES). A total of 44 fatty acids and two alcs. were found in members of the genus *Aeromonas*. All 90 strains contained 12:0, 13:0 iso, 14:0, 15:0 iso 3OH, 16:0, 16:1 .omega.7c, 17:0 iso, iso 17:1 .omega.9c, summed feature 3 (16:1 iso I and/or 14:0 3OH), and summed feature 7 (18:1 .omega.7c, 18:1 .omega.9t, and/or 18:1 .omega.12t), whereas all but one strain (99%) also contained 15:0 iso. Although the FAME profiles were very similar, minor quant. variations could be used to differentiate phenospecies and/or hybridization groups. A cluster anal. of the mean data revealed five FAME clusters, which were compared with phenotypic and genotypic groups identified in the genus *Aeromonas*. Hybridization groups that constituted the *Aeromonas hydrophila* complex, the *Aeromonas caviae* complex, and the *Aeromonas sobria* complex were basically grouped into distinct FAME clusters. The taxonomic positions of hybridization groups 7 and 11 in these clusters, however, remained unclear. All of our results were highly reproducible. A new database of *Aeromonas* FAME fingerprints was generated, and this database can be used for rapid identification of unknown aeromonads. Using a large set of well-characterized aeromonads, we demonstrated for the first time that gas-liq. chromatog. FAME anal. can be used to differentiate the majority of the phenospecies and/or hybridization groups in the genus *Aeromonas*.

L3 ANSWER 52 OF 53 CAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1994:647335 CAPLUS

DOCUMENT NUMBER: 121:247335

TITLE: Protein identification in DNA databases by peptide mass fingerprinting

AUTHOR(S): James, Peter; Quadroni, Manfredo; Carafoli, Ernesto; Gonnet, Gaston

CORPORATE SOURCE: Dep. Biology, Swiss Federal Institute Technology,  
Zurich, 8092, Switz.

SOURCE: Protein Science (1994), 3(8), 1347-50

CODEN: PRCIEI; ISSN: 0961-8368

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Proteins can be identified using a set of peptide fragment wts. produced by a specific digestion to search a protein database in which sequences have been replaced by fragment wts. calcd. for various cleavage methods. We present a method using multidimensional searches that greatly increases the confidence level for identification, allowing DNA sequence databases to be examd. This method provides a link between 2-dimensional gel electrophoresis protein databases and genome sequencing projects. Moreover, the increased confidence level allows unknown proteins to be matched to expressed sequence tags, potentially eliminating the need to obtain sequence information for cloning. Database searching from a mass profile is offered as a free service by an automatic server at the Swiss Federal Institute of Technol. (ETH), Zurich. For information, send an electronic message to the address [cborg@inf.ethz.ch](mailto:cborg@inf.ethz.ch) with the line: help

mass search, or help all.

L3 ANSWER 53 OF 53 CAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1993:599099 CAPLUS

DOCUMENT NUMBER: 119:199099

TITLE: Protein identification by mass profile fingerprinting

AUTHOR(S): James, Peter; Quadroni, Manfredo; Carafoli, Ernesto;  
Gonnet, Gaston

CORPORATE SOURCE: Protein Chem. Lab., Swiss Fed. Inst. Technol., Zurich,  
8092, Switz.

SOURCE: Biochemical and Biophysical Research Communications (1993), 195(1), 58-64

CODEN: BBRC A9; ISSN: 0006-291X

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The authors developed an algorithm for identifying proteins at the submicrogram level without sequence detn. by chem. degrdn. The protein, usually isolated by 1- or 2-dimensional gel electrophoresis, is digested by enzymic or chem. means and the masses of the resulting peptides are detd. by mass spectrometry. The resulting mass profile, i.e., the list of the mol. masses of peptides produced by the digestion, serves as a fingerprint which uniquely defines a particular protein. This fingerprint may be used to search the database of known sequences to find proteins with a similar profile. If the protein is not yet sequenced the profile can serve as a unique marker. This provides a rapid and sensitive link between genomic sequences and 2D gel electrophoresis mapping of cellular proteins.